High-Order Error Bounds for Markovian LSA with Richardson-Romberg Extrapolation

Ilya Levin[1], **Alexey Naumov**[1], **Sergey Samsonov**[1]
[1]HSE University
ivlevin@hse.ru, anaumov@hse.ru, svsamsonov@hse.ru

Abstract

In this paper, we study the bias and high-order error bounds of the Linear Stochastic Approximation (LSA) algorithm with Polyak-Ruppert (PR) averaging under Markovian noise. We focus on the version of the algorithm with constant step size and propose a novel decomposition of the bias via a linearization technique. We analyze the structure of the bias and show that the leading-order term is linear in the step size and cannot be eliminated by PR averaging. To address this, we apply the Richardson-Romberg (RR) extrapolation procedure, which effectively cancels the leading bias term. We derive high-order moment bounds for the RR iterates and show that the leading error term aligns with the asymptotically optimal covariance matrix of the vanilla averaged LSA iterates. We validate applicability of our findings for the temporal difference algorithm in reinforcement learning.

Extended version — https://arxiv.org/abs/2508.05570

1 Introduction

Stochastic approximation (SA) algorithms (Robbins and Monro 1951) play a foundational role in modern machine learning due to their various applications in reinforcement learning (Sutton and Barto 2018) and empirical risk minimization. In this paper, we consider the simplified setting of linear SA (LSA) algorithms, which estimate a solution of the linear system $\bar{\mathbf{A}}\theta^\star = \bar{\mathbf{b}}$. For a sequence of step sizes $\{\alpha_k\}_{k\in\mathbb{N}}$, a burn-in period $n_0 \in \mathbb{N}$, and an initialization $\theta_0 \in \mathbb{R}^d$, we consider the sequences of estimates $\{\theta_k\}_{k\in\mathbb{N}}$ and $\{\bar{\theta}_n\}_{n\geq n_0+1}$ given by

$$\begin{aligned} \theta_k &= \theta_{k-1} - \alpha_k\{\mathbf{A}(Z_k)\theta_{k-1} - \mathbf{b}(Z_k)\}\ , \quad k \geq 1, \\ \bar{\theta}_n &= (n-n_0)^{-1}\sum_{k=n_0}^{n-1}\theta_k\ , \quad n \geq n_0+1\ . \end{aligned} \tag{1}$$

Here, $\bar{\theta}_n$ corresponds to the Polyak-Ruppert averaged estimator (Ruppert 1988; Polyak and Juditsky 1992), a popular instrument for accelerating the convergence of stochastic approximation algorithms. In (1), $\{Z_k\}_{k\in\mathbb{N}}$ is a sequence of random variables taking values in some measurable space $(\mathsf{Z}, \mathcal{Z})$, and $\mathbf{A}(Z_k)$ and $\mathbf{b}(Z_k)$ are stochastic estimates of $\bar{\mathbf{A}}$ and $\bar{\mathbf{b}}$, respectively. In this paper, we focus on the setting where $\{Z_k\}_{k\in\mathbb{N}}$ is a Markov chain.

One of the key questions related to the recurrence (1) is the choice of step sizes $\{\alpha_k\}_{k\in\mathbb{N}}$. While the classical SA schemes (Robbins and Monro 1951; Polyak and Juditsky 1992) correspond to the setting of decreasing step sizes, a lot of recent contributions (Huo et al. 2024; Lauand and Meyn 2022) focus on the setting of constant step sizes $\alpha_k = \alpha > 0$. This setting is of particular interest because it enables geometrically fast forgetting of the initialization (Dieuleveut, Durmus, and Bach 2020) and is often easier to use in practice. At the same time, the solution of the SA problem obtained with a constant step size suffers from an inevitable *bias*, which arises in non-linear problems (Dieuleveut, Durmus, and Bach 2020) or even in linear SA (1) when the sequence of noise variables $\{Z_k\}_{k\in\mathbb{N}}$ forms a Markov chain, see e.g., (Lauand and Meyn 2022; Durmus et al. 2025; Huo, Chen, and Xie 2023a). This problem can be partially mitigated using the Richardson-Romberg (RR) extrapolation method. To formally define this method, we denote the LSA iterations (1) with a constant step size α and define the corresponding Polyak-Ruppert averaged iterates as

$$\begin{aligned} \theta_k^{(\alpha)} &= \theta_{k-1}^{(\alpha)} - \alpha\{\mathbf{A}(Z_k)\theta_{k-1}^{(\alpha)} - \mathbf{b}(Z_k)\}\ , \\ \bar{\theta}_n^{(\alpha)} &= (n-n_0)^{-1}\sum_{k=n_0}^{n-1}\theta_k^{(\alpha)}\ . \end{aligned} \tag{2}$$

The next steps of the Richardson-Romberg (RR) procedure rely on the fact that the bias of $\bar{\theta}_n^{(\alpha)}$ is linear in α and is of order $\mathcal{O}(\alpha)$, see e.g., (Huo, Chen, and Xie 2023a). To proceed further, a learner considers two sequences $\{\theta_k^{(\alpha)}, k \in \mathbb{N}\}$ and $\{\theta_k^{(2\alpha)}, k \in \mathbb{N}\}$ with the same noise sequence $\{Z_k\}_{k\in\mathbb{N}}$. Then for any $n \geq n_0+1$, one can set

$$\bar{\theta}_n^{(\alpha,\mathsf{RR})} = 2\bar{\theta}_n^{(\alpha)} - \bar{\theta}_n^{(2\alpha)}\ .$$

The non-asymptotic analysis of Richardson-Romberg extrapolation has recently attracted a lot of contributions in the context of linear SA (Huo, Chen, and Xie 2023a), stochastic gradient descent (SGD) (Durmus et al. 2016; Dieuleveut, Durmus, and Bach 2020), and non-linear SA problems (Huo et al. 2024; Allmeier and Gast 2024). At the same time, a large and relatively unexplored gap is related to the question of the optimality of the leading term of the error bounds for $\bar{\theta}_n^{(\alpha,\mathsf{RR})} - \theta^\star$. To properly define what "optimality" means in this context, note that in the context of linear SA problems with a decreasing step size (1), the sequence $\{\bar{\theta}_n\}_{n\in\mathbb{N}}$ is asymptotically normal under appropriate conditions on $\{\alpha_k\}_{k\in\mathbb{N}}$, that is

$$\sqrt{n}(\bar{\theta}_n - \theta^\star) \xrightarrow{d} \mathcal{N}(0, \Sigma_\infty), \quad n \to \infty .$$

The covariance matrix Σ_∞ here is known to be asymptotically optimal both in a sense of the Rao-Cramer lower bound and in a sense that it corresponds to the last iterate of the modified process $\tilde{\theta}_k$, which uses the optimal preconditioner matrix ($\bar{A}^{-1}$ in the context of linear SA). Details can be found in the papers (Polyak and Juditsky 1992; Fort, Gersende 2015). A precise expression for Σ_∞ is given later in the current paper, see (7). It is known for SGD methods with i.i.d. noise and averaging that the Richardson-Romberg estimator achieves mean-squared error bounds (MSE) with the leading term, which aligns with Σ_∞; that is,

$$\mathbb{E}^{1/2}[\|\bar{\theta}_n^{(\alpha,\mathsf{RR})} - \theta^\star\|^2] \le \frac{\sqrt{\operatorname{Tr}\Sigma_\infty}}{\sqrt{n}} + \mathcal{O}\left(\frac{1}{n^{1/2+\delta}}\right) ,$$

for some $\delta > 0$. This result is due to (Sheshukova et al. 2025). To the best of our knowledge, there is no result of this kind available for the setting of Markovian SA. In this paper, we aim to close this gap for the setting of linear SA, yet we expect that the developed method can be useful for a more general setting. The main contributions of this paper are as follows:

- We propose a novel technique to quantify the asymptotic bias of $\theta_n^{(\alpha)}$. Our approach considers the limiting distribution Π_α of the joint Markov chain $\{(\theta_k^{(\alpha)}, Z_{k+1})\}_{k\in\mathbb{N}}$ and analyzes the bias $\Pi_\alpha(\theta_0) - \theta^\star$. Then, we apply the linearization method for $\theta_k^{(\alpha)}$ from (Aguech, Moulines, and Priouret 2000). This allows us to study the limiting distribution of the components, whose average values are shown to be ordered by powers of α.
- We establish high-order moment error bounds for the Richardson-Romberg method, where the leading term aligns with the asymptotically optimal covariance Σ_∞. We analyze its dependence on the number of steps n, step size α, and the mixing time t_{mix}.

2 Related Work

The stochastic approximation scheme is widely studied for reinforcement learning (RL) (Sutton 1988; Sutton and Barto 2018). The well-known Temporal-Difference (TD) algorithm with linear function approximation (Bertsekas and Tsitsiklis 1996) can be represented as the LSA problem. Originally, this method was proposed in (Robbins and Monro 1951) with a diminishing step size. While asymptotic convergence results were first studied, non-asymptotic analysis later became of particular interest. For general SA, non-asymptotic bounds were investigated in (Moulines and Bach 2011; Gadat and Panloup 2023). For LSA with a constant step size, finite-time analysis was presented in (Mou et al. 2020, 2024; Durmus et al. 2025).

The bias and MSE for non-linear problems with i.i.d. noise have been studied for SGD in (Dieuleveut, Durmus, and Bach 2020; Yu et al. 2021; Sheshukova et al. 2025), and, recently, with both i.i.d. and Markovian noise in (Zhang and Xie 2024; Zhang et al. 2024; Huo et al. 2024; Allmeier and Gast 2024). Another source of bias arises under Markovian noise and cannot be eliminated using averaging, as shown in (Lauand and Meyn 2022, 2023). MSE bounds for Markovian LSA have been studied in several works, including (Srikant and Ying 2019; Mou et al. 2024; Durmus et al. 2025). In (Mou et al. 2024) and (Durmus et al. 2025), the authors derive the leading term, which aligns with the optimal covariance Σ_∞, but they do not eliminate the effect of the asymptotic bias.

Further, when studying Markovian LSA, in (Lauand and Meyn 2022) the authors address the problem of bias, which can't be eliminated using PR averaging. In the work (Lauand and Meyn 2023), the authors establish weak convergence of the Markov chain (θ_n, Z_{n+1}) and also provide a decomposition for the limiting covariance of the iterations. In our work, we establish a similar result in Theorem 1. The work (Lauand and Meyn 2024) extends results on bias and convergence of Polyak-Ruppert iterations to diminishing step sizes $\alpha_k = \alpha_0 k^{-\rho}$ with $\rho \in (0, 1/2)$.

The non-asymptotic analysis of Richardson-Romberg has been carried out in (Durmus et al. 2016; Huo et al. 2024; Sheshukova et al. 2025; Allmeier and Gast 2024) for general SA, with particular applications to SGD. Further, in (Huo, Chen, and Xie 2023a) and (Huo, Chen, and Xie 2023b), the authors derive bounds for the LSA problem. In the work (Huo et al. 2024), the authors establish a bias decomposition for general SA up to the linear term in the step size α and derive MSE bounds dependent on α and the mixing time. For LSA, (Huo, Chen, and Xie 2023a) extends this analysis by deriving a bias decomposition via an infinite series expansion in α and examining the MSE under the RR procedure, which eliminates arbitrary leading-order terms. Both works demonstrate that the RR technique accelerates convergence and maintains the proper scaling with the mixing time. However, neither work explicitly identifies the leading-term coefficient, and their results primarily address the improvement of higher-order terms in α. Additionally, (Huo, Chen, and Xie 2023a) imposes a restrictive reversibility assumption on the underlying Markov chain, limiting its applicability. Separately, (Huo, Chen, and Xie 2023b) explores the role of the RR procedure in statistical inference, particularly in constructing confidence intervals. Further, in (Zhang and Xie 2024; Kwon et al. 2025) authors consider the application of the RR procedure for Q-learning and two-timescale SA. A comparison of the bias decompositions known in the literature with our approach can be found in Section 4.

3 Notations

Consider a Polish space Z and a Markov kernel Q on $(\mathsf{Z}, \mathcal{Z})$ endowed with its Borel σ-field denoted by $\mathcal{Z}$ and let $(\mathsf{Z}^{\mathbb{N}}, \mathcal{Z}^{\otimes\mathbb{N}})$ be the corresponding canonical space. Consider a Markov kernel Q on $\mathsf{Z} \times \mathcal{Z}$ and denote by $\mathbb{P}_\xi$ and $\mathbb{E}_\xi$ the corresponding probability distribution and expectation with initial distribution ξ. Without loss of generality, assume that

$(Z_k)_{k\in\mathbb{N}}$ is the associated canonical process. By construction, for any $\mathsf{A} \in \mathcal{Z}$, $\mathbb{P}_\xi\left(Z_k \in \mathsf{A} \mid Z_{k-1}\right) = \mathrm{Q}(Z_{k-1}, \mathsf{A})$, $\mathbb{P}_\xi$-a.s. In the case $\xi = \delta_z$, $z \in \mathsf{Z}$, $\mathbb{P}_\xi$ and $\mathbb{E}_\xi$ are denoted by $\mathbb{P}_z$ and $\mathbb{E}_z$. Also, for any measurable space $(\mathsf{X}, \mathcal{G})$ with the signed measure μ, we define the total variation norm $\|\mu\|_{\mathsf{TV}} = |\mu|(\mathsf{X})$.

Let $(\mathsf{X}, \mathcal{G})$ be a complete separable metric space equipped with its Borel σ-algebra $\mathcal{G}$. We call $c : \mathsf{X} \times \mathsf{X} \to \mathbb{R}_+$ a distance-like function, if it is symmetric, lower semi-continuous and $c(x, y) = 0$ if and only if $x = y$, and there exists $q \in \mathbb{N}$ such that $d(x, y)^q \leq c(x, y)$. We denote by $\mathcal{H}(\xi, \xi')$ the set of couplings of probability measures ξ and ξ', that is, a set of probability measures on $(\mathsf{X} \times \mathsf{X}, \mathcal{G} \otimes \mathcal{G})$, such that for any $\Gamma \in \mathcal{H}(\xi, \xi')$ and any $A \in \mathcal{G}$ it holds $\Gamma(\mathsf{X} \times A) = \xi'(A)$ and $\Gamma(A \times \mathsf{X}) = \xi(A)$. We define the Wasserstein semimetric associated to the distance-like functiton $c^p(\cdot, \cdot)$, as

$$\mathbf{W}_{c,p}(\xi, \xi') = \inf_{\Gamma \in \mathcal{H}(\xi, \xi')} \int_{\mathsf{X}\times\mathsf{X}} c^p(x, x')\Gamma(dx, dx') \; . \quad (3)$$

We also denote $\mathbf{W}_c(\xi, \xi') := \mathbf{W}_{c,1}(\xi, \xi')$.

4 Bias of the LSA Iterates

In this section we aim to study the properties of the sequence $\theta_k^{(\alpha)}$ given by (2) based on theory of Markov chains. Using the definition (2) and some elementary algebra, we obtain

$$\theta_k^{(\alpha)} - \theta^\star = (\mathrm{I} - \alpha\mathbf{A}(Z_k))(\theta_{k-1}^{(\alpha)} - \theta^\star) - \alpha\varepsilon(Z_k) \; , \quad (4)$$

where we have set

$$\varepsilon(z) = \tilde{\mathbf{A}}(z)\theta^\star - \tilde{\mathbf{b}}(z) \; , \quad \tilde{\mathbf{A}}(z) = \mathbf{A}(z) - \bar{\mathbf{A}} \; , \quad (5)$$
$$\tilde{\mathbf{b}}(z) = \mathbf{b}(z) - \bar{\mathbf{b}} \; .$$

We consider the following assumptions on the noise variables $\{Z_k\}$:

UGE 1. *$\{Z_k\}_{k\in\mathbb{N}}$ is a Markov chain with the Markov kernel Q taking values in complete separable metric space $(\mathsf{Z}, \mathcal{Z})$. Moreover, Q admits π as an invariant distribution and is uniformly geometrically ergodic, that is, there exists $t_{\mathrm{mix}} \in \mathbb{N}^*$ such that for all $k \in \mathbb{N}^*$,*

$$\Delta(\mathrm{Q}^k) \leq (1/4)^{\lfloor k/t_{\mathrm{mix}} \rfloor} \; , \quad (6)$$

where $\Delta(\mathrm{Q}^k)$ is Dobrushin coefficient defined as

$$\Delta(\mathrm{Q}^k) = \sup_{z,z'\in\mathsf{Z}} (1/2)\|\mathrm{Q}^k(z, \cdot) - \mathrm{Q}^k(z', \cdot)\|_{\mathsf{TV}} \; .$$

Equivalently, there exist constants $\zeta > 0$ and $\rho \in (0, 1)$ such that for all $k \geq 1$,

$$\sup_{z\in\mathsf{Z}} \|\mathrm{Q}^k(z, \cdot) - \pi\|_{\mathsf{TV}} \leq \zeta\rho^k \; .$$

Here, t_{mix} is the mixing time of Q. **UGE** 1 implies, in particular, that π is the unique invariant distribution of Q. We also define the noise covariance matrix

$$\Sigma_\varepsilon^{(\mathrm{M})} = \mathbb{E}_\pi[\varepsilon(Z_0)\varepsilon(Z_0)^T] + 2\sum_{\ell=1}^{\infty} \mathbb{E}_\pi[\varepsilon(Z_0)\varepsilon(Z_\ell)^T] \; .$$

This covariance is limiting for the sum $n^{-1/2}\sum_{t=0}^{n-1}\varepsilon(Z_t)$, see (Douc et al. 2018)[Theorem 21.2.10]. Due to (Fort, Gersende 2015), the asymptotically optimal covariance matrix Σ_∞ is defined as

$$\Sigma_\infty = (\bar{\mathbf{A}})^{-1}\Sigma_\varepsilon^{(\mathrm{M})}(\bar{\mathbf{A}})^{-T} \; . \quad (7)$$

In the considered setting when $\{Z_k\}_{k\in\mathbb{N}}$ is a Markov chain, the sequence $\{\theta_k^{(\alpha)}\}$ given by (4), considered separately from $\{Z_k\}_{k\in\mathbb{N}}$, might fail to be a Markov chain. This is not the case in the setting when Z_k are i.i.d. random variables, see e.g. (Mou et al. 2020; Durmus et al. 2025). That is why, in the current paper we need to consider the joint process $(\theta_k^{(\alpha)}, Z_{k+1})$, which is a Markov chain with the kernel $\bar{\mathrm{P}}_\alpha$, specified below. For any measurable and bounded function $f : \mathbb{R}^d \times \mathsf{Z} \to \mathbb{R}_+$, $(\theta, z) \in \mathbb{R}^d \times \mathsf{Z}$, we define $\bar{\mathrm{P}}_\alpha$ as

$$\bar{\mathrm{P}}_\alpha f(\theta, z) = \int_{\mathsf{Z}} \mathrm{Q}(z, \mathrm{d}z') f(\mathbf{F}_{z'}(\theta), z') \; ,$$
$$\mathbf{F}_z(\theta) = (\mathrm{I} - \alpha\mathbf{A}(z))\theta + \alpha\mathbf{b}(z) \; .$$

Thus, our next aim is to perform a quantitative analysis of $\bar{\mathrm{P}}_\alpha$. In particular, we show below that under appropriate regularity conditions, $\bar{\mathrm{P}}_\alpha$ admits a unique invariant distribution Π_α. Specifically, we impose the following assumptions:

A1. $\mathrm{C}_{\mathbf{A}} = \sup_{z\in\mathsf{Z}}\|\mathbf{A}(z)\| \vee \sup_{z\in\mathsf{Z}}\|\tilde{\mathbf{A}}(z)\| < \infty$ *and the matrix $-\bar{\mathbf{A}}$ is Hurwitz.*

In particular, the condition that $-\bar{\mathbf{A}}$ is Hurwitz implies that the linear system $\bar{\mathbf{A}}\theta = \bar{\mathbf{b}}$ has a unique solution $\theta^\star$. We further require the following assumptions on the noise term $\varepsilon(z)$ and the stationary distribution π of the sequence $\{Z_k\}_{k\in\mathbb{N}^*}$:

A2. $\int_{\mathsf{Z}} \mathbf{A}(z)\mathrm{d}\pi(z) = \bar{\mathbf{A}}$ *and* $\int_{\mathsf{Z}} \mathbf{b}(z)\mathrm{d}\pi(z) = \bar{\mathbf{b}}$. *Moreover,* $\|\varepsilon\|_\infty = \sup_{z\in\mathsf{Z}}\|\varepsilon(z)\| < +\infty$.

Theorem 1. *Assume **A** 1, **A** 2, and **UGE** 1. Let $2 \leq p \leq q$. Then, for any $\alpha \in (0, (\alpha_{q,\infty}^{(M)} \wedge a^{-1})t_{\mathrm{mix}}^{-1})$, the Markov kernel $\bar{\mathrm{P}}_\alpha$ admits a unique invariant distribution Π_α, such that $\Pi_\alpha(\|\theta_0 - \theta^\star\|) < \infty$. Here $\alpha_{q,\infty}^{(\mathrm{M})}$ is a constant depending upon q and other problem characteristics, and is defined in* (30).

Proof sketch. We consider two noise sequences, $\{Z_n, n \in \mathbb{N}\}$ and $\{\tilde{Z}_n, n \in \mathbb{N}\}$, with a coupling time T. They evolve separately before time T and coincide afterwards. See more details on coupling construction in Appendix B.1. To prove the statement, we first establish the result on the contraction of the Wasserstein semimetric (3) with the cost function c_0, defined as

$$\mathsf{c}_0((\theta, z), (\theta', z')) = (\|\theta - \theta'\| + \mathbf{1}_{\{z\neq z'\}})$$
$$\times \left(1 + \|\theta - \theta^\star\| + \|\theta' - \theta^\star\|\right) \; ,$$

where $(\theta, z), (\theta', z') \in \mathbb{R}^d \times \mathsf{Z}$. To do that, we consider two coupled Markov chains $\{(\theta_k^{(\alpha)}, Z_{k+1}), k \geq 0\}$ and $\{(\tilde{\theta}_k^{(\alpha)}, \tilde{Z}_{k+1}), k \geq 0\}$, starting from (θ, z) and $(\tilde{\theta}, \tilde{z})$ respectively. For $n \geq 1$ and $\theta, \tilde{\theta} \in \mathbb{R}$, we define:

$$\theta_n^{(\alpha)} = \theta_{n-1}^{(\alpha)} - \alpha\{\mathbf{A}(Z_n)\theta_{n-1}^{(\alpha)} - \mathbf{b}(Z_n)\}, \quad \theta_0 = \theta \; ,$$
$$\tilde{\theta}_n^{(\alpha)} = \tilde{\theta}_{n-1}^{(\alpha)} - \alpha\{\mathbf{A}(\tilde{Z}_n)\tilde{\theta}_{n-1}^{(\alpha)} - \mathbf{b}(\tilde{Z}_n)\}, \quad \theta_0 = \tilde{\theta} \; .$$

Then, for any $z, z' \in \mathsf{Z}$, from the result in Proposition 4, we get:

$$\tilde{\mathbb{E}}_{z,\tilde{z}}[\mathsf{c}_0((\theta_n^{(\alpha)}, Z_n), (\tilde{\theta}_n^{(\alpha)}, \tilde{Z}_n))] \lesssim \rho_\alpha^n \mathsf{c}_0((z,\theta),(\tilde{z},\tilde{\theta})) \;, \quad (8)$$

where $\rho_\alpha = \mathrm{e}^{-\alpha a/24}$ and the expectation is taken over the coupling measure. Finally, the existence and uniqueness of the invariant measure Π_α follows from the contraction inequality (8) in conjunction with (Douc et al. 2018, Theorem 20.3.4). The detailed proof is provided in Appendix B.1. □

Our next goal is to quantify the bias

$$\Pi_\alpha[\theta_0] - \theta^\star \;.$$

Towards this aim, we consider the perturbation-expansion framework of (Aguech, Moulines, and Priouret 2000), see also (Durmus et al. 2025). We define the product of random matrices

$$\Gamma_{m:n}^{(\alpha)} = \prod_{i=m}^{n} (\mathrm{I} - \alpha \mathbf{A}(Z_i)) \;, \quad m \leq n \;, \quad (9)$$

with the convention, $\Gamma_{m:n}^{(\alpha)} = \mathrm{I}$ for $m > n$. Then we consider the decomposition of the error into the transient and fluctuation terms

$$\theta_n^{(\alpha)} - \theta^\star = \tilde{\theta}_n^{(\mathsf{tr})} + \tilde{\theta}_n^{(\mathsf{fl})} \;, \quad (10)$$

where

$$\tilde{\theta}_n^{(\mathsf{tr})} = \Gamma_{1:n}^{(\alpha)}\{\theta_0 - \theta^\star\} \;, \quad (11)$$

$$\tilde{\theta}_n^{(\mathsf{fl})} = -\alpha \sum_{j=1}^{n} \Gamma_{j+1:n}^{(\alpha)} \varepsilon(Z_j) \;.$$

Bounding the transient and fluctuation terms To bound the transient term, we apply the result on exponential stability of the random matrix product from (Durmus et al. 2025, Proposition 7). For the fluctuation term $\tilde{\theta}_n^{(\mathsf{fl})}$ we use the perturbation expansion technique formalized in (Aguech, Moulines, and Priouret 2000) and later applied to obtain the high-probability bounds in (Durmus et al. 2025). For this decomposition, we define for any $l \geq 0$ the vectors $\{J_n^{(l,\alpha)}, H_n^{(l,\alpha)}\}$ which can be computed from the recursion relations

$$J_n^{(0,\alpha)} = \left(\mathrm{I} - \alpha \bar{\mathbf{A}}\right) J_{n-1}^{(0,\alpha)} - \alpha \varepsilon(Z_n) \;, \quad (12)$$

$$H_n^{(0,\alpha)} = (\mathrm{I} - \alpha \mathbf{A}(Z_n))\, H_{n-1}^{(0,\alpha)} - \alpha \tilde{\mathbf{A}}(Z_n) J_{n-1}^{(0,\alpha)} \;, \quad (13)$$

where $J_0^{(0,\alpha)} = H_0^{(0,\alpha)} = 0$. It is easy to check that

$$\tilde{\theta}_n^{(\mathsf{fl})} = J_n^{(0,\alpha)} + H_n^{(0,\alpha)} \;.$$

Moreover, the term $H_n^{(0,\alpha)}$ can be further decomposed simialry to (12) - (13). Precisely, for any $L \in \mathbb{N}^*$ and $\ell \in \{1, \ldots, L\}$, we consider

$$J_n^{(\ell,\alpha)} = \left(\mathrm{I} - \alpha \bar{\mathbf{A}}\right) J_{n-1}^{(\ell,\alpha)} - \alpha \tilde{\mathbf{A}}(Z_n) J_{n-1}^{(\ell-1,\alpha)} \;, \quad (14)$$

and

$$H_n^{(\ell,\alpha)} = (\mathrm{I} - \alpha \mathbf{A}(Z_n))\, H_{n-1}^{(\ell,\alpha)} - \alpha \tilde{\mathbf{A}}(Z_n) J_{n-1}^{(\ell,\alpha)} \;, \quad (15)$$

where we set $J_0^{(l,\alpha)} = H_0^{(l,\alpha)} = 0$. It is easy to check that, in this setting,

$$H_n^{(l,\alpha)} = J_n^{(l+1,\alpha)} + H_n^{(l+1,\alpha)} \;,$$

and

$$\tilde{\theta}_n^{(\mathsf{fl})} = \sum_{\ell=0}^{L} J_n^{(0,\alpha)} + H_n^{(L,\alpha)} \;. \quad (16)$$

To analyze the bias $\Pi_\alpha[\theta_0] - \theta^\star$, we consider this expansion with $L = 2$. That is, combining (12), (13) and (14), we obtain the decomposition which is the cornerstone of our analysis:

$$\theta_n^{(\alpha)} - \theta^\star = \tilde{\theta}_n^{(\mathsf{tr})} + J_n^{(0,\alpha)} + J_n^{(1,\alpha)} + J_n^{(2,\alpha)} + H_n^{(2,\alpha)} \;. \quad (17)$$

Following the arguments in (Durmus et al. 2021), this decomposition can be used to obtain sharp bounds on the p-th moment of the final LSA iterate $\theta_n^{(\alpha)}$.

Bias expansion for LSA Similarly to (4), we can not consider the process $\{J_k^{(\ell,\alpha)}\}$ separately, as it might fail to be a Markov chain. Instead, we again consider the joint process

$$Y_t = (Z_{t+1}, J_t^{(0,\alpha)}, J_t^{(1,\alpha)}) \quad (18)$$

with the Markov kernel $\mathrm{Q}_{J^{(1)}}$, which can be defined formally in the similar way as $\bar{\mathrm{P}}_\alpha$. We need to refine our assumptions on the step-size compared to (31). More specifically, for any $2 \leq p < \infty$, we set

$$\alpha_{p,\infty}^{(\mathsf{b})} = \left(\alpha_{p(1+\log d),\infty}^{(\mathrm{M})} \wedge \frac{1}{1 + \mathrm{C}_{\mathbf{A}}} \wedge \frac{1}{ap}\right) t_{\mathrm{mix}}^{-1} \;, \quad (19)$$

where $\alpha_\infty^{(\mathrm{M})}$ is defined in (31). For ease of notation, we set $\alpha_\infty^{(\mathsf{b})} := \alpha_{2,\infty}^{(\mathsf{b})}$. Note that the established step size suggests to take smaller step sizes in order to control higher moments.

Proposition 1. *Assume **A** 1, **A** 2 and **UGE** 1. Let* $\alpha \in (0, \alpha_\infty^{(\mathsf{b})})$. *Then the process* $\{Y_t\}_{t\in\mathbb{N}}$ *is a Markov chain with a unique stationary distribution* $\Pi_{J^{(1)},\alpha}$.

Proof sketch. We consider the Markov chain $\{Y_t, t \geq 0\}$ with kernel $\mathrm{Q}_{J^{(1)}}$, where $Y_t = (Z_{t+1}, J_t^{(0,\alpha)}, J_t^{(1,\alpha)})$. Our approach involves analyzing the convergence of this Markov chain using the Wasserstein semimetric, defined in (3), with a properly chosen cost function. Denoting $Y = (z, J^{(0)}, J^{(1)})$ and $\tilde{Y} = (\tilde{z}, \tilde{J}^{(0)}, \tilde{J}^{(1)})$, where $Y, \tilde{Y} \in \mathsf{Z} \times \mathbb{R}^d \times \mathbb{R}^d$, we define the cost function as:

$$\begin{aligned} \mathsf{c}(Y, \tilde{Y}) = {} & \|J^{(0)} - \tilde{J}^{(0)}\| + \|J^{(1)} - \tilde{J}^{(1)}\| \\ & + (\|J^{(0)}\| + \|\tilde{J}^{(0)}\| \\ & + \|J^{(1)}\| + \|\tilde{J}^{(1)}\| + \sqrt{\alpha a}\|\varepsilon\|_\infty)\mathbf{1}_{\{z \neq \tilde{z}\}} \;. \end{aligned} \quad (20)$$

Note, the term $\sqrt{\alpha a}\|\varepsilon\|_\infty$ is introduced to account for the fluctuations of $J_n^{(0,\alpha)}$ and $J_n^{(1,\alpha)}$, whose magnitudes do not exceed the order of this term. Now, we introduce the result on the contraction of the Wasserstein semimetric for two coupled Markov chains $\{Y_t\}$ and $\{\tilde{Y}_t\}$ starting from different points. Choosing $J^{(0)}, \tilde{J}^{(0)}, J^{(1)}, \tilde{J}^{(1)} \in \mathbb{R}^d$ and $z, \tilde{z} \in$

Z, we denote $y = (z, J^{(0)}, J^{(1)})$ and $\tilde{y} = (\tilde{z}, \tilde{J}^{(0)}, \tilde{J}^{(1)})$ such that $y \neq \tilde{y}$. Then, by Lemma 3 with $p = 1$, for any $n \geq 1$, we have

$$\mathbf{W}_{\mathsf{c}}(\delta_y \mathrm{Q}^n_{J^{(1)}}, \delta_{\tilde{y}} \mathrm{Q}^n_{J^{(1)}}) \lesssim \rho^n_{1,\alpha} \sqrt{\log(1/\alpha a)}\, \mathsf{c}(y, \tilde{y}) \;, \quad (21)$$

where $\rho_{1,\alpha} = e^{-\alpha a/12}$. Thus, the existence of invariant distribution $\Pi_{J^{(1)},\alpha}$ directly follows from (21) and (Douc et al. 2018, Theorem 20.3.4); for more details, see Appendix B.3. □

We denote random variables

$$(Z_{\infty+1}, J^{(0,\alpha)}_{\infty}, J^{(1,\alpha)}_{\infty})$$

with distribution $\Pi_{J^{(1)},\alpha}$. Under stationary distribution, we have $\mathbb{E}[J^{(0,\alpha)}_{\infty}] = 0$. Consider now the component that corresponds to $J^{(1,\alpha)}_{\infty}$. The following proposition holds:

Proposition 2. *Assume* **A** *1,* **A** *2 and* **UGE** *1. Then for $\alpha \in (0, \alpha^{(\mathsf{b})}_{\infty})$, it holds that*

$$\lim_{n\to\infty} \mathbb{E}[J^{(1,\alpha)}_n] = \mathbb{E}[J^{(1,\alpha)}_{\infty}] = \alpha\Delta + R(\alpha) \;,$$

where $\Delta \in \mathbb{R}^d$ is defined as

$$\Delta = \bar{\mathbf{A}}^{-1} \sum_{k=1}^{\infty} \mathbb{E}[\tilde{\mathbf{A}}(Z_{\infty+k})\varepsilon(Z_{\infty})] \;,$$

and $R(\alpha)$ is a remainder term which can be bounded as

$$\|R(\alpha)\| \leq 12 \|\bar{\mathbf{A}}^{-1}\| \, \mathrm{C}^2_{\mathbf{A}} \, t^2_{\mathrm{mix}} \alpha^2 \|\varepsilon\|_{\infty} \;.$$

Corollary 1. *Under the setting of Proposition 2, we get the following expansion for the asymptotic bias*

$$\lim_{n\to\infty} \mathbb{E}[\theta_n] = \Pi_{\alpha}(\theta_0) = \theta^\star + \alpha\Delta + O(\alpha^{3/2}) \;. \quad (22)$$

Proof. From Proposition 8, we deduce that $\lim_{n\to\infty} \mathbb{E}\big[\|J^{(2,\alpha)}_n\|\big] \lesssim \alpha^{3/2}$ and $\lim_{n\to\infty} \mathbb{E}\big[\|H^{(2,\alpha)}_n\|\big] \lesssim \alpha^{3/2}$. This implies that the term $J^{(1,\alpha)}_n$ should be the leading term in the bias decomposition. Together with the analysis of $J^{(2,\alpha)}_n$, this confirms that $\{J^{(l+1,\alpha)}_n, l \geq 0\}$ provides the proper linearization of the bias in powers of α, giving rigorous justification for our decomposition approach. For the complete proof, we refer to Appendix B.5. □

Remark 1. *By sequentially analyzing the terms $\{J^{(k,\alpha)}_n, k \geq 2\}$ in the decomposition of θ_n, we can obtain the bias decomposition as a power series in α. Additionally, in Proposition 6, we show that*

$$\lim_{n\to\infty} \mathbb{E}[J^{(2,\alpha)}_n] = \alpha^2\Delta_2 + R_2(\alpha) \;,$$

where

$$\Delta_2 = -\sum_{k=1}^{\infty}\sum_{i=0}^{\infty} \mathbb{E}[\tilde{\mathbf{A}}(Z_{\infty+k+i+1})\tilde{\mathbf{A}}(Z_{\infty+i+1})\varepsilon(Z_{\infty})]$$

and $\|R_2(\alpha)\| \lesssim \alpha^{5/2}$. Furthermore, we expect that the remainder term $H^{(2,\alpha)}_n = J^{(3,\alpha)}_n + H^{(3,\alpha)}_n$ in the bias decomposition (17) *is of order $\mathcal{O}(\alpha^2)$. To establish this result, we can use a technique similar to the one used for the p-th moment of $J^{(2,\alpha)}_n$ in Proposition 8. This suggests that $\mathbb{E}^{1/p}[\|J^{(3,\alpha)}_n\|^p] \lesssim \alpha^2$. Therefore, we conclude that the rate $\mathcal{O}(\alpha^2)$ could be achieved in the remainder term of* (22)*.*

Discussion Our coefficient Δ in the linear term matches the representation derived in (Lauand and Meyn 2023, Theorem 2.5), but that work does not analyze MSE with reduced bias. To observe the next result, we define an adjoint kernel Q^* such that for the invariant measure π, we have $\pi \otimes \mathrm{Q}(A \times B) = \pi \otimes \mathrm{Q}^*(B \times A)$. Additionally, we define the independent kernel Π such that for any $z \in \mathsf{Z}$ and $A \in \mathcal{Z}$, $\Pi(z, A) = \pi(A)$. Under these notations, the authors in (Huo, Chen, and Xie 2023a) considered the bias expansion arising from the Neumann series for the operator $(\mathrm{I} - \mathrm{Q}^* + \Pi)^{-1}(\mathrm{Q}^* - \Pi)$. Furthermore, adapting the proof of Proposition 2, our result can be reformulated in terms of Q^*. This representation is less desirable because it requires reversibility of the Markov kernel Q, as discussed in (Huo, Chen, and Xie 2023a).

5 Analysis of Richardson-Romberg Procedure

A natural way to reduce the bias in (22) is to use the Richardson-Romberg extrapolation (Hildebrand 1987)

$$\bar{\theta}^{(\alpha,\mathsf{RR})}_n = 2\bar{\theta}^{(\alpha)}_n - \bar{\theta}^{(2\alpha)}_n \;. \quad (23)$$

After this procedure the remainder term in the bias has order $O(\alpha^{3/2})$. Before the main theorem of this section, we establish our key technical results. For that, we consider another Markov chain $\{V_t\}_{t\in\mathbb{N}}$ with kernel Q_J, where we set $V_t = (J_t, Z_{t+1})$. In fact, it is closely related to the one described in (18) and also converges geometrically fast to the unique stationary distribution as stated by Corollary 2.

Corollary 2. *Assume* **A** *1,* **A** *2 and* **UGE** *1. Let $\alpha \in (0, \alpha^{(\mathsf{b})}_{\infty})$. Then the process $\{V_t\}_{t\in\mathbb{N}}$ is a Markov chain with a unique stationary distribution $\Pi_{J,\alpha}$.*

Proof sketch. We define the cost function $\mathsf{c}_J : \mathbb{R}^d \times \mathsf{Z} \times \mathbb{R}^d \times \mathsf{Z} \to \mathbb{R}_+$, as:

$$\mathsf{c}_J((J,z),(\tilde{J},\tilde{z})) = \|J - \tilde{J}\| + (\|J\| + \|\tilde{J}\| + \sqrt{\alpha a}\|\varepsilon\|_{\infty})\mathbf{1}_{\{z\neq\tilde{z}\}} \;.$$

The result on the contraction of the Wasserstein semimetric with cost function c_J for $\{V_t\}_{t\in\mathbb{N}}$ can be also obtained independently using the technique from Proposition 1. However, we derive a weaker result directly from Proposition 1, showing that $\mathbf{W}_{\mathsf{c}_J,p}(\delta_y \mathrm{Q}^n_J, \delta_{\tilde{y}} \mathrm{Q}^n_J) \leq \mathbf{W}_{\mathsf{c},p}(\delta_y \mathrm{Q}^n_{J^{(1)}}, \delta_{\tilde{y}} \mathrm{Q}^n_{J^{(1)}})$. Hence, using the similar arguments, we conclude that the Markov chain $\{V_t, t \geq 0\}$ admits invariant distribution $\Pi_{J,\alpha}$. The detailed proof can be found in Appendix C. □

Note that the invariant distribution $\Pi_{J,\alpha}$ coincides with the distribution of $(J^{(0,\alpha)}_{\infty}, Z_{\infty+1})$. For any $J \in \mathbb{R}^d, z \in \mathsf{Z}$, we define:

$$\bar{\psi}(J,z) = \psi(J,z) - \mathbb{E}_{\pi_J}[\psi_0],$$
$$\bar{\psi}_t = \bar{\psi}(J^{(0,\alpha)}_t, Z_{t+1}) \;.$$

The cost functions c_J and $\mathsf{c}_{J^{(1)}}$ are designed such that the function $\psi(J,z) = \tilde{\mathbf{A}}(z)J$ for $J \in \mathbb{R}^d, z \in \mathsf{Z}$ is Lipschitz, specifically:

$$\|\psi(J,z) - \psi(\tilde{J},\tilde{z})\| \leq 2\,\mathrm{C}_{\mathbf{A}}\,\mathsf{c}_J((J,z),(\tilde{J},\tilde{z})) \;.$$

This Lipschitz property is necessary for our analysis of Theorem 2. The following result concerns the magnitude of $\sum_{t=n_0}^{n-1} \psi_t$, which appears in the decomposition (26). It has a non-zero bias, and thus, a direct estimation leads to non-optimal behavior. However, after centering, the result in Proposition 3 suggests that it can be estimated effectively. This provides a theoretical justification for the numerical experiments presented in Section 6.

Proposition 3. *Assume **A** 1, **A** 2, and **UGE** 1. Then for any probability measure ξ on $\mathbb{R}^d \times \mathsf{Z}$, $2 \le p < \infty$ and $\alpha \in (0, \alpha_{p,\infty}^{(\mathsf{b})})$, we get*

$$\mathbb{E}_\xi^{1/p}[\| \sum_{t=0}^{n-1} \bar{\psi}_t \|^p] \le \mathrm{c}_{W,1}^{(2)} p^{3/2} (\alpha n)^{1/2} + \mathrm{c}_{W,2}^{(2)} p^3 \alpha^{-1/2} \sqrt{\log(1/\alpha a)} \;,$$

where the constants $\mathrm{c}_{W,1}^{(2)}, \mathrm{c}_{W,2}^{(2)}$ are defined in the supplement paper, see (68).

Now, we conclude the result on p-th moment for error of the RR iteration (23).

Theorem 2. *Assume **A** 1, **A** 2 and **UGE** 1. Fix $2 \le p < \infty$, then for any $n \ge t_{\mathrm{mix}}$, $\alpha \in (0, \alpha_{p,\infty}^{(\mathsf{b})})$ and initial probability measure ξ on $(\mathsf{Z}, \mathcal{Z})$, we have*

$$\begin{aligned} &\mathbb{E}_\xi^{1/p}[\| \bar{\mathbf{A}}(\bar{\theta}_n^{(\alpha,\mathsf{RR})} - \theta^\star) \|^p] \qquad (24) \\ &\le 2\, \mathrm{C}_{\mathsf{Rm},1} \{ \operatorname{Tr} \Sigma_\varepsilon^{(\mathrm{M})} \}^{1/2} p^{1/2} n^{-1/2} + R_{n,p,\alpha}^{(\mathsf{fl})} \\ &+ R_{n,p,\alpha}^{(\mathsf{tr})} \| \theta_0 - \theta^\star \| \exp\{ -\alpha a n / 24 \} \;, \end{aligned}$$

where $R_{n,p,\alpha}^{(\mathsf{tr})}$, $R_{n,p,\alpha}^{(\mathsf{fl})}$ are provided in (25)*, and $\mathrm{C}_{\mathsf{Rm},1} = 60\mathrm{e}$ is obtained from the Rosenthal inequality(see Theorem 3).*

The quantities $R_{n,p,\alpha}^{(\mathsf{tr})}$ and $R_{n,p,\alpha}^{(\mathsf{fl})}$ correspond to the fluctuation and transient terms in the error decomposition. We set them as follows

$$\begin{aligned} R_{n,p,\alpha}^{(\mathsf{fl})} &\lesssim p n^{-3/4} \qquad (25) \\ &+ (p^{3/2} (\alpha n)^{-1/2} + \alpha^{1/2}) p^{3/2} n^{-1/2} \\ &+ p^{7/2} \alpha^{3/2} \log^{3/2}(1/\alpha a) \;, \\ R_{n,p,\alpha}^{(\mathsf{tr})} &\lesssim (\alpha n)^{-1} \;, \end{aligned}$$

Here $\lesssim$ stands for the inequality up a constant which doesn't depend on p, n and α. Precise expressions for the terms $R_{n,p,\alpha}^{(\mathsf{fl})}$ and $R_{n,p,\alpha}^{(\mathsf{tr})}$ are given in the supplement paper, see equation (87).

Proof sketch of Theorem 2. Using (1) and the definition of the noise term $\varepsilon(\cdot)$ in (5), we can write the decomposition for the Richardson-Romberg iterations

$$\begin{aligned} &\bar{\mathbf{A}}(\bar{\theta}_n^{(\alpha,\mathsf{RR})} - \theta^\star) \qquad (26) \\ &= \{ 2\alpha(n - n_0) \}^{-1} (4\theta_{n_0}^{(\alpha)} - \theta_{n_0}^{(2\alpha)} - (4\theta_n^{(\alpha)} - \theta_n^{(2\alpha)})) \\ &+ \{ n - n_0 \}^{-1} \sum_{t=n_0}^{n-1} \{ e(\theta_t^{(2\alpha)}, Z_{t+1}) - 2e(\theta_t^{(\alpha)}, Z_{t+1}) \} \;. \end{aligned}$$

The leading term can be bounded using the result for the p-th moment of the last iteration in Lemma 7. The last term can be further decomposed using

$$\textstyle\sum_{t=n_0}^{n-1} e\left(\theta_t^{(\alpha)}, Z_{t+1} \right) = E_n^{(\mathsf{tr},\alpha)} + E_n^{(\mathsf{fl},\alpha)} \;, \qquad (27)$$

where we have set

$$\begin{aligned} E_n^{(\mathsf{tr},\alpha)} &= \textstyle\sum_{t=n_0}^{n-1} \tilde{\mathbf{A}}(Z_{t+1}) \Gamma_{1:t}^{(\alpha)} \{ \theta_0 - \theta^\star \} \;, \\ E_n^{(\mathsf{fl},\alpha)} &= \textstyle\sum_{t=n_0}^{n-1} \varepsilon(Z_{t+1}) \\ &+ \sum_{\ell=0}^{2} \sum_{t=n_0}^{n-1} \tilde{\mathbf{A}}(Z_{t+1}) J_t^{(\ell,\alpha)} + \sum_{t=n_0}^{n-1} \tilde{\mathbf{A}}(Z_{t+1}) H_t^{(2,\alpha)} \;. \end{aligned}$$

The first term in $E_n^{(\mathsf{fl},\alpha)}$ is linear statistics of Markov chain $\{Z_k, k \in \mathbb{N}\}$. Therefore, we can bound it using the version of Rosenthal inequality for Markov chains from (Durmus et al. 2023). For the term involving $J_t^{(0,\alpha)}$, we employ the expansion from (88), yielding a centered random variable component plus a bias term. This decomposition allows direct application of the inequality in Proposition 3 to the sum of centered random variables, which yields the bound $\mathcal{O}((\alpha/n)^{1/2} + \alpha^{-1/2} n^{-1})$. Combining this with the result from Proposition 2, we conclude that the remaining term is $\mathcal{O}(\alpha^2)$.

Then, we apply Proposition 8 to control the statistic $\sum_{t=n_0}^{n-1} \tilde{\mathbf{A}}(Z_{t+1}) J_t^{(1,\alpha)}$, which we express in terms of $J_n^{(2,\alpha)}$ via the expansion in (14). For the analogous term involving $H_n^{(2,\alpha)}$, we establish the required bound in Proposition 9. The detailed proof can be found in Appendix D.1. □

Note that the bound in Proposition 3 motivates the choice $\alpha = \mathcal{O}(n^{-1/2})$, aligning with the rate observed in the i.i.d case. Optimization over α gives us the following high-probability bound. Also, the term with p^3 could be slightly improved to p^2 through a more accurate analysis of the Lemma 5. Additionally, following the discussion in Remark 1, we expect that the remainder term $\mathcal{O}(\alpha^{3/2})$ in Theorem 2 could be improved to $\mathcal{O}(\alpha^2)$, though this would require a technically complicated analysis of $J_n^{(3,\alpha)}$. Using Markov's inequality, we derive the following high-probability bound.

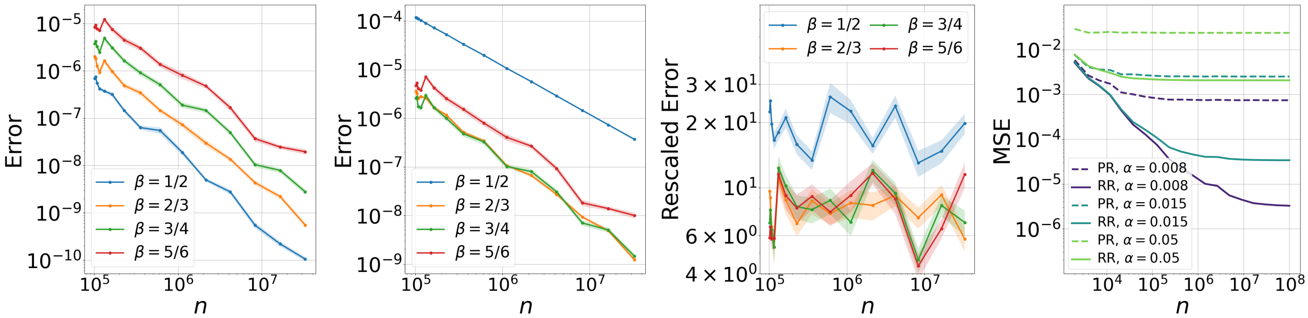
Corollary 3. *Assume **A** 1, **A** 2 and **UGE** 1. For $2 \le p < \infty$ and any $n \ge t_{\mathrm{mix}}$, we consider the step size*

$$\alpha(n, d, t_{\mathrm{mix}}, p) = \alpha_{p,\infty}^{(\mathsf{b})} n^{-1/2} \;. \qquad (28)$$

Substituting (28) *into* (24) *with $p = \ln(3\mathrm{e}/\delta)$, it holds with probability at least $1 - \delta$, that*

$$\begin{aligned} \| \bar{\mathbf{A}}(\bar{\theta}_n^{(\alpha,\mathsf{RR})} - \theta^\star) \| &\lesssim \sqrt{\operatorname{Tr} \Sigma_\varepsilon^{(\mathrm{M})} \log(1/\delta)} n^{-1/2} \\ &+ (1 + \log^{3/2}(n) \log^{5/2}(1/\delta)) \log(1/\delta) n^{-3/4} \\ &+ n^{-1/2} \log(1/\delta) \| \theta_0 - \theta^\star \| \exp\left\{ -\alpha_{1+\log d,\infty}^{(\mathsf{b})} n^{1/2} \right\} \;. \end{aligned}$$

Discussion Our analysis establishes high-order moment bounds and, as a consequence, high-probability bounds for RR iterations in Markovian LSA. Moreover, the leading



(a) (b) (c) (d)

Figure 1: In subfigures (a), (b) and (c), we consider the step size $\alpha = n^{-\beta}$. Subfigure (a): error for RR iterations (23). Subfigure (b): error for PR iterations (2). Subfigure (c): error for RR, multiplied by a factor $n^{2-\beta}$, which corresponds to the leading term of (25) after substituting α. Subfigure (d): MSE of Polyak-Ruppert and Richardson-Romberg iterations for different step sizes α.

term in (24) scales with $\{\operatorname{Tr}\Sigma_\varepsilon^{(\mathrm{M})}\}^{1/2}$, which is known to be locally asymptotically minimax optimal for the Polyak-Ruppert iterates (Mou et al. 2024) and aligns with the CLT covariance matrix Σ_∞(see (7)). In (Dieuleveut, Durmus, and Bach 2020), the authors study the bias and MSE for SGD with i.i.d noise, and propose the Richardson-Romberg extrapolation to reduce this bias. However, they only consider MSE bounds and do not obtain the proper factor for the leading term. In the Markovian LSA literature, the authors similarly consider only MSE and do not explicitly emphasize the leading term (Huo et al. 2024; Huo, Chen, and Xie 2023a,b; Zhang and Xie 2024). The closest result, (Sheshukova et al. 2025), shows high-order bounds with the leading term properly aligned with the optimal covariance, but in this work, the authors consider general SA with i.i.d. noise, the analysis of which differs significantly from our case.

6 Experiments

In this section, we aim to demonstrate the effect of reduced bias achieved through Richardson-Romberg extrapolation and to validate the accuracy of the bound obtained in Theorem 2. For this purpose, we adopt an example introduced in (Lauand and Meyn 2024). More precisely, we consider the Markovian noise $\{Z_k, k \geq 1\}$ on the space $\mathsf{Z} = \{0, 1\}$ with transition matrix $P = \begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix}$ and $a \in (0,1)$. For any $z \in \{0,1\}$, we consider the noisy observations

$$\mathbf{A}(z) = z \cdot A^{(1)} + (1-z) \cdot A^{(0)},$$
$$\mathbf{b}(z) = z \cdot b^{(1)} + (1-z) \cdot b^{(0)} \ ,$$

where we set

$$A^{(0)} = -2\begin{pmatrix} -2 & 0 \\ 1 & -2 \end{pmatrix}, \quad b^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
$$A^{(1)} = -2\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad b^{(1)} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

Hence, we have $\bar{\mathbf{A}} = \mathrm{I}$ and $\bar{\mathbf{b}} = (1/2)b^{(1)}$. In the following experiments, we set $a = 0.3$ and ran $N_{\mathsf{traj}} = 400$ trajectories from $\theta_0 = \theta^\star$ following (2).

Figure 1d illustrates the significant reduction in bias achieved by the Richardson-Romberg scheme, estimating $\mathbb{E}[\|\bar{\theta}_n^{(\alpha)} - \theta^\star\|^2]$ and $\mathbb{E}[\|\bar{\theta}_n^{(\alpha,\mathsf{RR})} - \theta^\star\|^2]$. These results justify that, after a few iterations, the error of the RR procedure starts to decrease faster than for PR averaging. Additionally, in Figure 1, we show that the resulting dependence on α and n in the bounds (25) is tight. To achieve this, for different sample size n we select different step sizes of the form $\alpha = n^{-\beta}$ for $\beta \in [1/2, 1)$, substitute these into (25), and compute the scaling of the term $R^{(\mathsf{fl})}_{n,p,\alpha}$ w.r.t. n. For $\beta \geq 1/2$, with mentioned choice of α, $R^{(\mathsf{fl})}_{n,p,\alpha}$ scales as $n^{\beta-2}$.

To verify numerically this rate, we consider the following procedure. We approximate the terms $\mathbb{E}[\|\bar{\theta}_n^{(\alpha)} - \theta^\star + (1/n)\sum_{k=1}^n \varepsilon(Z_k)\|^2]$ for PR averaging, and

$$\Delta_n^{(RR)} = \mathbb{E}[\|\bar{\theta}_n^{(\alpha,\mathsf{RR})} - \theta^\star + (1/n)\sum_{k=1}^n \varepsilon(Z_k)\|^2]$$

for Richardson-Romberg iterations. The moments of the latter term should scale with $n^{\beta-2}$. We verify this effect numerically setting $\alpha = n^{-\beta}$ for $\beta \in \{1/2, 2/3, 3/4, 5/6\}$ and providing the plots for $\Delta_n^{(RR)}$ and $n^{2-\beta}\Delta_n^{(RR)}$ in Figure 1a and Figure 1c, respectively. Additionally, in Figure 1a and Figure 1b, we compare the error for different choices of step α. We can see that the step $\alpha = n^{-1/2}$ gives the smallest error for Richardson-Romberg iterations, while for Polyak-Ruppert averaging this choice of step introduces a large bias in the error.

7 Conclusion

We studied the high-order error bounds for Richardson-Romberg extrapolation in the setting of Markovian linear stochastic approximation. By applying the novel technique for bias characterization, we were able to obtain the leading term which aligns with the asymptotically optimal covariance matrix Σ_∞. For further work, we consider the generalization of the obtained results to the setting of non-linear Markovian SA and SA with state-dependent noise.

Acknowledgments

The authors are grateful to Eric Moulines for valuable discussions and feedback. The work was supported by the grant for research centers in the field of AI provided by the Ministry of Economic Development of the Russian Federation in accordance with the agreement 000000C313925P4E0002 and the agreement with HSE University № 139-15-2025-009. This research was supported in part through computational resources of HPC facilities at HSE University (Kostenetskiy, Chulkevich, and Kozyrev 2021).

References

Aguech, R.; Moulines, E.; and Priouret, P. 2000. On a perturbation approach for the analysis of stochastic tracking algorithms. *SIAM Journal on Control and Optimization*, 39(3): 872–899.

Allmeier, S.; and Gast, N. 2024. Computing the Bias of Constant-step Stochastic Approximation with Markovian Noise. In Globerson, A.; Mackey, L.; Belgrave, D.; Fan, A.; Paquet, U.; Tomczak, J.; and Zhang, C., eds., *Advances in Neural Information Processing Systems*, volume 37, 137873–137902. Curran Associates, Inc.

Bertsekas, D.; and Tsitsiklis, J. N. 1996. *Neuro-dynamic programming*. Athena Scientific.

Dieuleveut, A.; Durmus, A.; and Bach, F. 2020. Bridging the gap between constant step size stochastic gradient descent and Markov chains. *The Annals of Statistics*, 48(3): 1348 – 1382.

Douc, R.; Moulines, E.; Priouret, P.; and Soulier, P. 2018. *Markov chains*. Springer Series in Operations Research and Financial Engineering. Springer. ISBN 978-3-319-97703-4.

Durmus, A.; Moulines, E.; Naumov, A.; and Samsonov, S. 2025. Finite-time high-probability bounds for Polyak–Ruppert averaged iterates of linear stochastic approximation. *Mathematics of Operations Research*, 50(2): 935–964.

Durmus, A.; Moulines, E.; Naumov, A.; Samsonov, S.; Scaman, K.; and Wai, H.-T. 2021. Tight High Probability Bounds for Linear Stochastic Approximation with Fixed Stepsize. In Ranzato, M.; Beygelzimer, A.; Nguyen, K.; Liang, P. S.; Vaughan, J. W.; and Dauphin, Y., eds., *Advances in Neural Information Processing Systems*, volume 34, 30063–30074. Curran Associates, Inc.

Durmus, A.; Moulines, E.; Naumov, A.; Samsonov, S.; and Sheshukova, M. 2023. Rosenthal-type inequalities for linear statistics of Markov chains. *arXiv preprint arXiv:2303.05838*.

Durmus, A.; Simsekli, U.; Moulines, E.; Badeau, R.; and RICHARD, G. 2016. Stochastic Gradient Richardson-Romberg Markov Chain Monte Carlo. In Lee, D.; Sugiyama, M.; Luxburg, U.; Guyon, I.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc.

Fort, Gersende. 2015. Central limit theorems for stochastic approximation with controlled Markov chain dynamics. *ESAIM: PS*, 19: 60–80.

Gadat, S.; and Panloup, F. 2023. Optimal non-asymptotic analysis of the Ruppert–Polyak averaging stochastic algorithm. *Stochastic Processes and their Applications*, 156: 312–348.

Hildebrand, F. 1987. *Introduction to Numerical Analysis*. Dover books on advanced mathematics. Dover Publications. ISBN 9780486653631.

Huo, D.; Chen, Y.; and Xie, Q. 2023a. Bias and extrapolation in Markovian linear stochastic approximation with constant stepsizes. In *Abstract Proceedings of the 2023 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, 81–82.

Huo, D.; Chen, Y.; and Xie, Q. 2023b. Effectiveness of Constant Stepsize in Markovian LSA and Statistical Inference. arXiv:2312.10894.

Huo, D. L.; Zhang, Y.; Chen, Y.; and Xie, Q. 2024. The collusion of memory and nonlinearity in stochastic approximation with constant stepsize. *Advances in Neural Information Processing Systems*, 37: 21699–21762.

Kostenetskiy, P.; Chulkevich, R.; and Kozyrev, V. 2021. HPC resources of the higher school of economics. In *Journal of Physics: Conference Series*, volume 1740, 012050. IOP Publishing.

Kwon, J.; Dotson, L.; Chen, Y.; and Xie, Q. 2025. Two-Timescale Linear Stochastic Approximation: Constant Stepsizes Go a Long Way. In *The 28th International Conference on Artificial Intelligence and Statistics*.

Lauand, C. K.; and Meyn, S. 2022. Bias in Stochastic Approximation Cannot Be Eliminated With Averaging. In *2022 58th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 1–4.

Lauand, C. K.; and Meyn, S. 2023. The curse of memory in stochastic approximation. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, 7803–7809. IEEE.

Lauand, C. K.; and Meyn, S. 2024. Revisiting step-size assumptions in stochastic approximation. *arXiv preprint arXiv:2405.17834*.

Mou, W.; Li, C. J.; Wainwright, M. J.; Bartlett, P. L.; and Jordan, M. I. 2020. On linear stochastic approximation: Fine-grained Polyak-Ruppert and non-asymptotic concentration. In *Conference on Learning Theory*, 2947–2997. PMLR.

Mou, W.; Pananjady, A.; Wainwright, M.; and Bartlett, P. 2024. Optimal and instance-dependent guarantees for Markovian linear stochastic approximation. *Mathematical Statistics and Learning*, 7.

Moulines, E.; and Bach, F. 2011. Non-asymptotic analysis of stochastic approximation algorithms for machine learning. *Advances in neural information processing systems*, 24: 451–459.

Pinelis, I. 1994. Optimum Bounds for the Distributions of Martingales in Banach Spaces. *The Annals of Probability*, 22(4): 1679 – 1706.

Polyak, B. T.; and Juditsky, A. B. 1992. Acceleration of stochastic approximation by averaging. *SIAM journal on control and optimization*, 30(4): 838–855.

Rio, E. 2017. *Asymptotic Theory of Weakly Dependent Random Processes*, volume 80. Springer.

Robbins, H.; and Monro, S. 1951. A stochastic approximation method. *The annals of mathematical statistics*, 400–407.

Ruppert, D. 1988. Efficient estimations from a slowly convergent Robbins-Monro process. Technical report, Cornell University Operations Research and Industrial Engineering.

Sheshukova, M.; Belomestny, D.; Durmus, A. O.; Moulines, E.; Naumov, A.; and Samsonov, S. 2025. Nonasymptotic Analysis of Stochastic Gradient Descent with the Richardson–Romberg Extrapolation. In *The Thirteenth International Conference on Learning Representations*.

Srikant, R.; and Ying, L. 2019. Finite-time error bounds for linear stochastic approximation and TD learning. In *Conference on Learning Theory*, 2803–2830. PMLR.

Sutton, R. S. 1988. Learning to predict by the methods of temporal differences. *Machine Learning*, 3(1): 9–44.

Sutton, R. S.; and Barto, A. G. 2018. *Reinforcement Learning: An Introduction*. The MIT Press, second edition.

Villani, C. 2009. *Optimal transport : old and new*. Grundlehren der mathematischen Wissenschaften. Berlin: Springer. ISBN 978-3-540-71049-3.

Yu, L.; Balasubramanian, K.; Volgushev, S.; and Erdogdu, M. A. 2021. An Analysis of Constant Step Size SGD in the Non-convex Regime: Asymptotic Normality and Bias. In Ranzato, M.; Beygelzimer, A.; Dauphin, Y.; Liang, P.; and Vaughan, J. W., eds., *Advances in Neural Information Processing Systems*, volume 34, 4234–4248. Curran Associates, Inc.

Zhang, Y.; Huo, D.; Chen, Y.; and Xie, Q. 2024. Prelimit coupling and steady-state convergence of constant-stepsize nonsmooth contractive sa. In *Abstracts of the 2024 ACM SIGMETRICS/IFIP PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems*, 35–36.

Zhang, Y.; and Xie, Q. 2024. Constant stepsize q-learning: Distributional convergence, bias and extrapolation. *arXiv preprint arXiv:2401.13884*.